

CLAIMS

What is claimed is:

1. A system for providing Java-implemented Application Servers to a plurality of clients, the system comprising:

5 a computing system having a plurality of processing elements, each element configured such that, despite a failure of one processing element, the remaining processing elements continue to function;

a plurality of Java-implemented Application Servers, at least one Server assigned to execute on one or more processing elements; and

10 a distributor module that is configured to capture connection requests from a client on a port, to select one of the plurality of Application Servers to communicate with the requesting client during the connection and to assign the connection request to the selected one of the plurality of Application Servers, such that, after the assignment, the selected Application Server communicates directly with the client.

15 2. A system for providing Java-implemented Application Servers to a plurality of clients as recited in claim 1, further comprising:

a monitoring module that is configured to restart the distributor and any Application server that fails.

20 3. A system for providing Java-implemented Application Servers to a plurality of clients as recited in claim 1,

wherein the distributor module is configured to select one of the Application Servers by:
determining whether there is an Application Server supporting fewer connections
25 than the other Application Servers,

if there is such an Application Server, sending to the Application Server a message that identifies the client making the connection request,

if there is no such Application Server, placing the client connection request on a waiting list until such an Application Server is available, and

when such an Application Server becomes available, sending to the Application Server a message that identifies the client making the connection request.

4. A system for providing Java-implemented Application Servers to a plurality of clients as recited in claim 3, wherein determining whether there is an Application Server supporting fewer connections than the other Application Servers is performed by executing a weighted round robin procedure.

5. A system for providing Java-implemented Application Servers to a plurality of clients as recited in claim 1, wherein the distributor module is further configured to:

maintain an information structure that includes a number indicating how many clients have a connection with each Application Server,
receive a message from an Application Server that a connection between the Application Server and a client is closed, and
update the number of clients with a connection to the Application Server that sent the message.

6. A method for providing Java-implemented Application Servers to a plurality of clients, the method comprising:

providing a plurality of processing elements, each element configured such that, despite a failure of one processing element, the remaining processing elements continue to function;
causing a plurality of Java-implemented Application Servers to execute on the processing elements, at least one Server assigned to execute on one or more of the processing elements;
receiving incoming client connection requests at at least one port managed by a distributor module;
selecting one of the Application Servers to communicate with the client during the connection; and
assigning, by the distributor module, the connection request to the selected Application Server such that, after the assignment, the selected Application Server communicates with the client directly.

7. A method for providing Java-implemented Application Servers to a plurality of clients as recited in claim 5, further comprising the steps of:

monitoring the operation of the Application Servers and the distributor module for failures; and

if a failure occurs in the distributor module or one of the Application Servers, restarting the distributor module or Application Server.

8. A method for providing Java-implemented Application Servers to a plurality of clients as recited in claim 5, wherein the step of selection one of the Application Servers includes the steps of:

determining whether there is an Application Server supporting fewer connections than the other Application Servers;

if there is such an Application Server, sending to the Application Server a message that identifies the client making the connection request;

if there is no such Application Server, placing the client connection request on a waiting list until such an Application Server is available; and

when such an Application Server becomes available, sending to the Application Server a message that identifies the client making the connection request.

9. A method for providing Java-implemented Application Servers to a plurality of clients as recited in claim 5,

wherein the distributor module maintains an information structure that includes a number indicating how many clients have a connection with each Application Server; and

further comprising the steps of:

receiving a message from an Application Server that a connection between the Application Server and a client is closed; and

updating the number of clients with a connection to the Application Server that sent the message.